

CLAIMS:

1. A method for error detection within text transcribed from a first speech signal by an automatic speech-to-text transcription system, comprising synthesizing a second speech signal from the transcribed text, providing first and second speech signal outputs for a comparison between first and second speech signals for an identification of potential errors in the text.
2. The method according to claim 1, wherein the speed and/or the volume of the second speech signal matches the speed and/or the volume of the first speech signal.
3. The method according to claim 1 or 2, wherein a set of filter functions is applied to the first speech signal to approximate the spectrum of the first speech signal to the spectrum of the second speech signal.
4. The method according to any one of the claims 1 to 3, wherein the second speech signal is generated by applying an inverse speech transcription process, generating a feature vector sequence from the text, using (a) statistical models of the speech-to-text transcription system and (b) a state sequence obtained in the process of transcription of the text from the first speech signal.
5. The method according to any one of the claims 1 to 4, wherein a comparison signal is generated by subtracting or superimposing first and second speech signals.

6. The method according to claim 5, wherein the comparison signal is provided acoustically and/or visually.
7. The method according to claim 5 or 6, wherein an error indication is outputted when the amplitude of the comparison signal is beyond a predefined range.
8. The method according to claim 7, wherein the error indication is outputted visually within the transcribed text on a graphical user interface.
9. The method according to any one of the claims 5 to 8, further comprising a pattern recognition of the comparison signal in order to identify a pre-trained pattern of the comparison signal being indicative of a type of error in the text.
10. The method according to claim 9, wherein a correction suggestion is provided with a detected type of error in the generated text.
11. An error detection system for a speech-to-text transcription system providing a transcribed text (412) from a first speech signal (400), the error detection system comprising:
 - means for synthesizing a second speech signal (416) from the transcribed text (412),
 - means for providing first (400, 418) and second (416) speech signals for comparison between first and second speech signals for an identification of potential errors in the text (412).
12. The detection system according to claim 11, wherein a comparison signal is generated by means of subtracting or superimposing first (400, 418) and second (416) speech signals.

13. The detection system according to claim 11 or 12, wherein the first (400, 418) and second (416) speech signal and/or the comparison signal is provided acoustically or visually for error detection purpose.

14. The detection system according to claim 12 or 13, wherein an error indication is outputted when the comparison signal is beyond a predefined range.

15. The detection system according to any one of the claims 12 to 14, wherein a distinct pattern in the comparison signal is assigned to a certain type of error in the transcribed text (412) and a correction suggestion being provided with a detected type of error in the transcribed text.

16. A computer program product for error detection for a speech-to-text transcription system providing a transcribed text from a first speech signal, the computer program product comprising program means for:

- synthesizing a second speech signal from the transcribed text,
- matching speed and/or volume of the second speech signal to the speed and/or and volume of the first speech signal,
- providing first and second speech signal outputs for a comparison between first and second speech signals.

17. The computer program product according to claim 16, the computer program product comprising means for generating a comparison signal by means of subtracting or superimposing first and second speech signals.

18. The computer program product according to claim 16 or 17, the computer program product comprising means for providing the first and second speech signals and/or the comparison signal acoustically or visually for error detection purpose.

19. The computer program product according to claim 17 or 18, the computer program product comprising means for outputting an error indication when the comparison signal is beyond a predefined range.

20. The computer program product according to any one of the claims 17 to 19, the computer program product comprising means for assigning a distinct pattern in the comparison signal to a certain type of error in the transcribed text and providing a correction suggestion with a detected type of error in the transcribed text.